

The Anarchist Library  
Anti-Copyright



# Linguistics and Brain Science

Noam Chomsky

2000

Noam Chomsky  
Linguistics and Brain Science  
2000

Retrieved on 19<sup>th</sup> June 2021 from [www.chomsky.info](http://www.chomsky.info)  
Published in *A. Marantz, Y. Miyashita and W. O'Neil (eds.) Image, Language and Brain*. pp. 13–28

**[theanarchistlibrary.org](http://theanarchistlibrary.org)**

In the past half century, there has been intensive and often highly productive inquiry into the brain, behavior, and cognitive faculties of many organisms. The goal that has aroused the most enthusiasm is also likely to be the most remote, probably by orders of magnitude: an understanding of the human brain and human higher mental faculties, their nature, and the ways they enter into action and interaction.

From the outset, there has been no shortage of optimistic forecasts, even declarations by distinguished researchers that the mind-body problem has been solved by advances in computation, or that everything is essentially understood apart from the “hard problem” of consciousness. Such conclusions surely do not withstand analysis. To an objective outside observer — say, a scientist from Mars — the optimism too might seem rather strange, since there is also no shortage of much simpler problems that are poorly understood, or not at all.

Despite much important progress in many areas, and justified excitement about the prospects opened by newer technologies, I think that a degree of skepticism is warranted, and that it is wise

to be cautious in assessing what we know and what we might realistically hope to learn.

The optimism of the early postwar period had many sources, some of them a matter of social history, I believe. But it also had roots in the sciences, in particular, in successful integration of parts of biology within the core natural sciences. That suggested to many people that science might be approaching a kind of “last frontier,” the mind and the brain, which should fall within our intellectual grasp in due course, as was soon to happen with DNA.

Quite commonly, these investigations have adopted the thesis that “Things mental, indeed minds, are emergent properties of brains,” while recognizing that “these emergences are not regarded as irreducible but are produced by principles that control the interactions between lower level events — principles we do not yet understand.” The last phrase reflects the optimism that has been a persistent theme throughout this period, rightly or wrongly.

I am quoting a distinguished neuroscientist, Vernon Mountcastle of the Johns Hopkins University Institute of Mind/Brain. Mountcastle is introducing a volume of essays published by the American Academy of Arts and Sciences, with contributions by leading researchers, who review the achievements of the past half century in understanding the brain and its functions (“The Brain” 1998). The thesis on emergence is widely accepted in the field, often considered a distinctive contribution of the current era. In the last few years, the thesis has repeatedly been presented as an “astounding hypothesis,” “the bold assertion that mental phenomena are entirely natural and caused by the neurophysiological activities of the brain” and “that capacities of the human mind are in fact capacities of the human brain.” The thesis has also been offered as a “radical new idea” in the philosophy of mind that may at last put to rest Cartesian dualism, some believe, while others express doubt that the apparent chasm between body and mind can really be bridged.

Within the brain and cognitive sciences, many would endorse the position expressed by Harvard evolutionary biologist E.

- Griffin, D. R. 1994. “Animal Communication as Evidence of Animal Mentality.” In D. C. Gajdusek and G. M. McKhann, eds., *Evolution and Neurology of Language: Discussions in Neuroscience X*, 1±2.
- Grodzinsky, Y. 1990. *Theoretical Perspectives on Language Deficits*. Cambridge, Mass.: MIT Press.
- Grodzinsky, Y., and L. Finkel. 1998. “The Neurology of Empty Categories: Aphasics’ Failure to Detect Ungrammaticality.” *Journal of Cognitive Neuroscience* 10(2): 281±292.
- Hawkins, R. D., and E. R. Kandel. 1984. “Is There a Cell-Biological Alphabet for Simple Forms of Learning?” *Psychological Review* 91: 376±391.
- Holton, G. 1996. “On the Art of Scientific Imagination.” *Daedalus*, Spring 183±208.
- Hume, David. [1778] 1983. *History of England*. Vol. 6, chap. 71. Indianapolis: Liberty Fund.
- Jenkins, L. 2000. *Biolinguistics*. Cambridge, England: Cambridge University Press.
- Koyré, A. 1957. *From the Closed World to the Infinite Universe*. Baltimore: Johns Hopkins University Press.
- Lange, Friedrich A. 1925. *The History of Materialism*. London: Kegan Paul.
- Russell, B. 1929. *The Analysis of Matter*. Leipzig: Teubner.
- Schofeld, Robert E. 1970. *Mechanism and Materialism: British Natural Philosophy in an Age of Reason*. Princeton: Princeton University Press.
- Thackray, A. 1970. *Atoms and Powers*. Cambridge, Mass.: Harvard University Press.

## References

- Block, N. 1990. "The Computer Model of the Mind." In D. N. Osherson and Edward E. Smith, eds., *An Invitation to Cognitive Science*, vol. 3: *Thinking*. Cambridge, Mass.: MIT Press.
- "The Brain." *Daedalus*, Spring 1998 (special issue).
- Brock, William H. 1992. *The Norton History of Chemistry*. New York: Norton.
- Caramazza, A. 1997. "Brain and Language." In M. S. Gazzaniga, *Conversations in the Cognitive Neurosciences*. Cambridge, Mass.: MIT Press.
- Chomsky, N. 1975. *Reflections on Language*. New York: Pantheon. Reprint. New York: New Press, 1998.
- Chomsky, N. 2000. "Minimalist Inquiries: The Framework." In R. Martin, D. Michaels, and J. Uriagereka, eds., *Step by Step: Essays on Minimalist Syntax in Honor of Howard Lasnik*. Cambridge, Mass.: MIT Press.
- Deacon, T. W. 1997. *The Symbolic Species: The Co-Evolution of Language and the Brain*. New York: Norton.
- Fodor, J. A. 1983. *The Modularity of Mind*. Cambridge, Mass.: MIT Press.
- Gallistel, C. R. 1997. "Neurons and Memory." In M. S. Gazzaniga, *Conversations in the Cognitive Neurosciences*. Cambridge, Mass.: MIT Press.
- Gallistel, C. R., ed. 1990a. "Animal Cognition." *Cognition* 37 (special issue), 1±2.
- Gallistel, C. R. 1990b. "Representations in Animal Cognition: An Introduction." In C. R. Gallistel, ed., "Animal Cognition." *Cognition* 37 (special issue), 1±22.
- Gazzaniga, M. S. 1997. *Conversations in the Cognitive Neurosciences*. Cambridge, Mass.: MIT Press.
- Gould, J. L. 1990. "Honey Bee Cognition." In C. R. Gallistel, ed., "Animal Cognition." *Cognition* 37 (special issue), 83±104.
- O. Wilson in the same American Academy issue on the brain: "Researchers now speak confidently of a coming solution to the brain-mind problem," presumably along the lines of Mountcastle's thesis on emergence. One contributor, the eminent neurobiologist Semir Zeki, suggests that the brain sciences can even confidently anticipate addressing the creative arts, thus incorporating the outer limits of human achievement within the neurosciences. He also observes that the ability to recognize "a continuous vertical line is a mystery that neurology has not yet solved"; perhaps the word yet is a bit more realistic here.

As far as I am aware, the neural basis for the remarkable behavior of bees also remains a mystery. This behavior includes what appear to be impressive cognitive feats and also some of the few known analogues to distinctive properties of human language, notably the regular reliance on "displaced reference" — communication about objects not in the sensory field (Griffin 1994). The prospects for vastly more complex organisms seem considerably more remote.

Whatever one may speculate about current prospects, it is worth bearing in mind that the leading thesis about minds as emergent properties of brains is far from novel. It revives eighteenth-century proposals put forth for compelling reasons, by, among others, the famous English scientist Joseph Priestley, and before him, the French physician Julien Offray de la Mettrie. As Priestley formulated the thesis, "The powers of sensation or perception and thought" are properties of "a certain organized system of matter." Properties "termed mental are the result [of the] organical structure" of the brain and "the human nervous system" generally.

In other words, "Things mental, indeed minds, are emergent properties of brains" (Mountcastle). Priestley of course could not say how this emergence takes place, and we are not much better off after 200 years.

The reasons for the eighteenth-century conclusions about emergence were indeed compelling. I think the brain and cognitive sciences can learn some useful lessons from the rise of the emergence

thesis 200 years ago, and from the ways the sciences have developed since, right up to mid-twentieth century, when the assimilation of parts of biology to chemistry took place. The debates of the early part of this century about atoms, molecules, chemical structures and reactions, and related matters are strikingly similar to current controversies about mind and brain. I would like to digress for a moment on these topics — instructive and pertinent ones, I think.

The reasoning that led to the eighteenth-century emergence thesis was straightforward. The modern scientific revolution was inspired by the “mechanical philosophy,” the idea that the world is a great machine that could in principle be constructed by a master artisan and that is therefore intelligible to us, in a very direct sense. The world is a complex version of the clocks and other intricate automata that fascinated the seventeenth and eighteenth centuries, much as computers have provided a stimulus to thought and imagination in recent years — the change of artifacts has limited consequences for the basic issues, as Alan Turing demonstrated sixty years ago.

In that context, Descartes had been able to formulate a relatively clear mind-body problem: it arose because he observed phenomena that, he plausibly argued, could not be accounted for in terms of automata. He was proven wrong, for reasons he could never have guessed: nothing can be accounted for within the mechanical philosophy, even the simplest terrestrial and planetary motion. Newton established, to his great dismay, that “a purely materialistic or mechanistic physics ... is impossible” (Koyré 1957:210).

Newton was bitterly criticized by leading scientists of his day for reverting to the mysticism from which we were at last to be liberated by the scientific revolution. He was condemned for reintroducing “occult qualities” that are no different from the mysterious “sympathies” and “antipathies” of the neoscholastic Aristotelian physicists, which were much ridiculed. Newton agreed. He regarded his discoveries as an utter “absurdity,” and for the rest

infinitival Tense-headed phrases without complementizer or verbal phrases without external argument (Chomsky 2000).

It is impossible to spell out the details and the empirical basis here, but the categories are clearly defined, and there is evidence that they have a special role with regard to sound, meaning, and intricate syntactic properties, including the systems of uninterpretable elements, dislocation, and the derivational interpretation of the recursive function. It would be extremely interesting to see if the conclusions could be tested by online studies of language use, or from other approaches.

To the extent that the strong minimalist thesis holds, interface conditions assume renewed importance. They can no longer simply be taken for granted in some in-explicit way, as in most empirical work on language. Their precise nature becomes a primary object of investigation-in linguistics, in the brain sciences, in fact from every point of view.

Exactly how the story unfolds from here depends on the actual facts of the matter.

At the level of language and mind, there is a good deal to say, but this is not the place. Again, I think it makes sense to think of this level of inquiry as in principle similar to chemistry early in the twentieth century: in principle that is, not in terms of the depth and richness of the “bodies of doctrine” established.

A primary goal is to bring the bodies of doctrine concerning language into closer relation with those emerging from the brain sciences and other perspectives. We may anticipate that richer bodies of doctrine will interact, setting significant conditions from one level of analysis for another, perhaps ultimately converging in true unification. But we should not mistake truisms for substantive theses, and there is no place for dogmatism as to how the issues might move toward resolution. We know far too little for that, and the history of modern science teaches us lessons that I think should not be ignored.

That is, language is a kind of optimal solution to the minimal conditions it must meet to be usable at all. This strong minimalist thesis, as it is sometimes called, is highly controversial, and should be: it would be quite surprising if something like that turned out to be true. I think the research program stimulated by this thesis is promising. It has already yielded some interesting and surprising results, which may have suggestive implications for the inquiry into language and the brain. This thesis brings to prominence an apparent property of language that I already mentioned, and that might prove fundamental: the significance of semantically uninterpretable morphological features, and their special role in language variety and function, including the dislocation property.

Other consequences also suggest research directions that might be feasible and productive. One major question of linguistic research, from every perspective, is what George Miller years ago called *chunking*: what are the units that constitute expressions, for storage of information, and for access in production, perception, retrieval, and other operations? Some are reasonably clear: something like syllables, words, larger phrases of various kinds. Others that seem crucial are harder to detect in the stream of speech: phonological and morphological elements, dislocation structures, and semantically relevant configurations that may be scarcely reflected in the sound of an expression, sometimes not at all, and in this sense are “abstract.” That is, these elements are really present in the internal computation, but with only indirect effects, if any, on the phonetic output.

Very recent work pursuing minimalist theses suggests that two types of abstract phrases are implicated in a special way in linguistic processes. The two types are the closest syntactic analogues to full propositions, in the semantic sense. In more technical terms, these are clauses with tense/event structure as well as force-mood indicators, and verbal phrases with a full argument structure: full CPs and verbal phrases with an external argument, but not finite or

of his life sought some way around them: he kept searching for a “certain most subtle spirit which pervades and lies hid in all gross bodies,” and would account for motion, interaction, electrical attraction and repulsion, properties of light, sensation, and the ways in which “members of animal bodies move at the command of the will” — comparable mysteries, he felt.

Similar efforts continued for centuries, but always in vain. The absurdity was real, and simply had to be accepted. In a sense it was overcome in this century, but only by introducing what Newton and his contemporaries would have regarded as even greater absurdities. We are left with the “admission into the body of science of incomprehensible and inexplicable ‘facts’ imposed upon us by empiricism” (Koyré 1957:272).

Well before Priestley, David Hume wrote that “Newton seemed to draw off the veil from some of the mysteries of nature,” but “he shewed at the same time the imperfections of the mechanical philosophy; and thereby restored [Nature’s] ultimate secrets to that obscurity, in which they ever did and ever will remain” (Hume [1778] 1983:542). The world is simply not comprehensible to human intelligence, at least in the ways that early modern science had hoped and expected. In his classic study of the history of materialism, Friedrich Lange observes that their expectations and goals were abandoned, and we gradually “accustomed ourselves to the abstract notion of forces, or rather to a notion hovering in a mystic obscurity between abstraction and concrete comprehension.” Lange describes this as a “turning-point” in the history of materialism that removes the surviving remnants of the doctrine far from those of the “genuine Materialists” of the seventeenth century, and deprives them of much significance (Lange 1925:308).

The turning point also led gradually to a much weaker concept of intelligibility than the one that inspired the modern scientific revolution: intelligibility of theories, not of the world — a considerable difference, which may well bring into operation different faculties of mind, a topic some day for cognitive science, perhaps.

A few years after writing the introduction to the English translation of Lange's history, Bertrand Russell illustrated the distinction with an example reinvented recently and now a centerpiece of debates over consciousness. Russell pointed out that "a man who can see knows things which a blind man cannot know; but a blind man can know the whole of physics," so "the knowledge which other men have and he has not is not part of physics" (Russell 1929:389). Russell is referring to the "qualitative knowledge which we possess concerning mental events," which might not simply be a matter of conscious awareness, as the phenomenon of blindsight suggests. Some leading animal researchers hold that something similar may be true of bees (Griffin 1994). Russell's own conclusion is that the natural sciences seek "to discover the causal skeleton of the world," and can aim no higher than that. "Physics studies percepts only in their cognitive aspect; their other aspects lie outside its purview" (Russell 1929:391±392).

These issues are now very much alive, but let us put them aside and return to the intellectual crisis of eighteenth-century science.

One consequence was that the concept of "body" disappeared. There is just the world, with its many aspects: mechanical, chemical, electromagnetic, optical, mental — aspects that we may hope to unify somehow, but how no one knows. We can speak of "the physical world," if we like, but for emphasis, without implying that there is some other world — rather the way we speak of the "real truth," without meaning that there is some other kind of truth. The world has occult properties, which we try to comprehend as best we can, with our highly specific forms of intelligence, which may leave much of nature a mystery, at least if we ourselves are part of the biological world, not angels. There is no longer a "mind-body problem," because there is no useful notion of "body," of the "material" or "physical" world. The terms simply indicate what is more or less understood and assimilable in some manner to core physics, whatever that turns out to be. For individual psychology, the emergence hypothesis of contemporary neuroscience becomes a truism:

ing the central role in determining both the functioning and the superficial variety of language: inflectional morphology that lacks semantic interpretation. This narrow subcomponent may also be what is involved in the ubiquitous and rather surprising "dislocation" property of human language: the fact that phrases are pronounced in one position in a sentence, but understood as if they were in a different position, where their semantic role would be transparent.

Here there is some convergence with other approaches, including work by Alfonso Caramazza and others. These investigators have found dissociation of inflectional morphology from other linguistic processes in aphasia, and have produced some intriguing results that suggest that dislocation too may be dissociated (Caramazza 1997). A result of particular interest for the study of language is the distinction that Grodzinsky and Finkel report between dislocation of phrasal categories and of lexical categories (Grodzinsky 1990; Grodzinsky and Finkel 1998). That result would tend to confirm some recent ideas about distinctions of basic semantic, phonological, and syntactic properties of these two types of dislocation: head movement and XP-movement in technical terms.

Other recent linguistic work has led to a sharper focus on the "interface" relations between extralinguistic systems and the cognitive system of language—that is, the recursive procedure that generates expressions. The extralinguistic systems include sensorimotor and conceptual systems, which have their own properties independent of the language faculty. These systems establish what we might call "minimal design specifications" for the language faculty. To be usable at all, a language must be "legible" at the interface: the expressions it generates must consist of properties that can be interpreted by these external systems.

One thesis, which seems to me much more plausible than anyone could have guessed a few years ago, is that these minimal design specifications are also maximal conditions in nontrivial respects.

patterns mask principles of a different character that cannot be detected directly in arrangement of phenomena.

But filling in the huge gaps and finding the real principles and generalizations is only part of the problem. It is also necessary to account for the fact that all children acquire their languages: their own private languages, of course, from this point of view, just as their visual systems are their own, not a target they are attempting to reach or a community possession or some extrahuman organism that coevolved with them.

It quickly became clear that the two basic goals are in conflict. To describe the state attained, it seemed necessary to postulate a rich and complex system of rules, specific to the language and even specific to particular grammatical constructions: relative clauses in Japanese, verb phrases in Swahili, and so on. But the most elementary observations about acquisition of language showed that that cannot be even close to accurate. The child has insufficient (or no) evidence for elementary properties of language that were discovered, so it must be that they reflect the initial state of the language faculty, which provides the basic framework for languages, allowing only the kinds of marginal variation that experience could determine.

The tension between these two goals set the immediate research agenda forty years ago. The obvious approach was to try to abstract general properties of the complex states attained, attribute them to the initial state, and show that the residue is indeed simple enough to be acquired with available experience. Many such efforts more or less crystallized fifteen to twenty years ago in what is sometimes called the *principles-and-parameters* approach. The basic principles of language are properties of the initial state; the parameters can vary in limited ways and are set by experience.

To a large extent, the parameters furthermore seem to be lexical, in fact properties of a small subcomponent of the lexicon, particularly inflectional morphology. Some recent work suggests that an even smaller subpart of inflectional morphology may be play-

there is no coherent alternative, with the abandonment of materialism in any significant sense of the concept.

Of course, that leaves all empirical problems unsolved, including the question of how bees find a flower after watching the “waggle dance,” and how they know not even to leave the hive if the directions lead to the middle of a lake, it has been reported (Gould 1990). Also included are questions about the relation between the principles of human language and properties of cells. Included as well are the much more far-reaching problems that troubled Descartes and Newton about the “commands of the will,” including the normal use of language — innovative, appropriate, and coherent, but apparently uncaused. It is useful to remember that these problems underlie Descartes’s two-substance theory, which was put to rest by Newton, who showed that one of the two substances does not exist: namely body.

How do we address the real problems? I know of no better advice than the recommendations of the eighteenth-century English chemist Joseph Black: “Chemical affinity must be accepted as a first principle, which we cannot explain any more than Newton could explain gravitation, and let us defer accounting for the laws of affinity until we have established such a body of doctrine as Newton has established concerning the laws of gravitation” (Black, quoted in Schofeld 1970:226). That is pretty much what happened. Chemistry proceeded to establish a rich body of doctrine, “its triumphs ... built on no reductionist foundation but rather achieved in isolation from the newly emerging science of physics” (Thackray 1970). That continued until recently. What was finally achieved by Linus Pauling sixty years ago was unification, not reduction. Russell’s observation in 1929 that chemical laws “cannot at present be reduced to physical laws” turns out to have been misleading, in an important way (Russell 1929). Physics had to undergo fundamental changes, mainly in the 1920s, in order to be unified with basic chemistry, departing even more radically from commonsense notions of “the physical.” Physics had to “free itself” from “intuitive pictures” and

give up the hope of “visualizing the world,” as Heisenberg put it (quoted in Holton 1996:191), another long leap away from intelligibility in the sense of the scientific revolution of the seventeenth century, which brought about the “first cognitive revolution” as well.

The unification of biology and chemistry a few years later can be misleading. That was genuine reduction, but to a newly created physical chemistry; some of the same people were involved, notably Pauling. True reduction is not so common in the history of science, and need not be assumed automatically to be a model for what will happen in the future.

Prior to the unification of chemistry and physics in the 1930s, it was commonly argued by distinguished scientists, including Nobel Prize winners in chemistry, that chemistry is just a calculating device, a way to organize results about chemical reactions, sometimes to predict them. Chemistry is not about anything real. The reason was that no one knew how to reduce it to physics. That failure was later understood: reduction was impossible, until physics underwent a radical revolution. It is now clear — or should be clear — that the debates about the reality of chemistry were based on fundamental misunderstanding. Chemistry was “real” and “about the world” in the only sense of these concepts that we have: it was part of the best conception of how the world works that human intelligence had been able to contrive. It is impossible to do better than that.

The debates about chemistry a few years ago are in many ways echoed in the philosophy of mind and the cognitive sciences today — and theoretical chemistry, of course, is hard science, merging indistinguishably with core physics. It is not at the periphery of scientific understanding, like the brain and cognitive sciences, which are trying to study systems vastly more complex. I think these recent debates about chemistry, and their surprising outcome, may be instructive for the brain and cognitive sciences. We should follow Joseph Black’s good advice and try to construct “bodies of doctrine”

unclear. Gallistel introduced a compendium of review articles on the topic a few years ago by arguing that representations play a key role in animal behavior and cognition. Here *representation* is to be understood in the mathematical sense of isomorphism: a one-one relation between mind/brain processes and “an aspect of the environment to which these processes adapt the animal’s behavior”—for example, when an ant represents the corpse of a conspecific by its odor (Gallistel 1990b:2).

The results are extremely interesting, but it is not clear that they offer useful analogues for human conceptual representation, specifically, for what is called *phonetic* or *semantic representation*. They do not seem to provide a useful approach to the relation of phonology to motions of molecules, and research does not follow this course. Personally, I think the picture is more misleading than helpful on the meaning side of language, contrary to most contemporary work about meaning and reference.

Here particularly, I think we can learn a good deal from work on these topics in the early modern period, now mostly forgotten. When we turn to the organization and generation of representations, analogies break down very quickly beyond the most superficial level.

The “biolinguistic” approach is at the core of the modern study of language, at least as I understand it. The program was formulated with relative clarity about forty years ago. As soon as the first attempts were made to develop recursive procedures to characterize linguistic expressions, it instantly became clear that little was known, even about well-studied languages. Existing dictionaries and grammars, however extensive, provide little more than hints and a few generalizations. They tacitly rely on the unanalyzed “intelligence of the reader” to fill in the rest, which is just about everything. Furthermore the generalizations are often misleading or worse, because they are limited to observed phenomena and their apparent structural arrangements -morphological paradigms, for example. As has been discovered everywhere in the sciences, these



Ordinary ways of talking about language reinforce such conceptions. Thus we say that a child is learning English but has not yet reached the goal. What the child has acquired is not a language at all: we have no name for whatever it is that a four-year-old has acquired. The child has a “partial, and partially erroneous, grasp” of English. So does everyone, in fact.

Learning is an achievement. The learner has a goal, a target: you aim for the goal and if you have not reached it, you have not yet learned, though you may be on the way. Formal learning theory adopts a similar picture: it asks about the conditions that must be satisfied for the learner to reach the target, which is set independently. It also takes the “language” to be a set of sentences, not the recursive procedure for generating expressions in the sense of the empirical study of language (often called the *internalized grammar*, a usage that has sometimes been misleading). In English, unlike similar languages, one also speaks of “knowing a language.” That usage has led to the conclusion that some cognitive relation holds between the person and the language, which is therefore outside the person: we do not know a state of our brains.

None of this has any biological interpretation. Furthermore, much of it seems to me resistant to any explicit and coherent interpretation. That is no problem for ordinary language, of course. But there is no reason to suppose that common usage of such terms as *language* or *learning* (or *belief* or numerous others like them), or others belonging to similar semantic fields in other linguistic systems, will find any place in attempts to understand the aspects of the world to which they pertain. Likewise, no one expects the commonsense terms *energy* or *liquid* or *life* to play a role in the sciences, beyond a rudimentary level. The issues are much the same.

There have been important results in the study of animal behavior and communication in a variety of species, generally in abstraction from the cellular level. How much such work advances us toward an understanding of human higher mental faculties seems

in whatever terms we can, unshackled by commonsense intuitions about how the world must be — we know that it is not that way — and untroubled by the fact that we may have to “defer accounting for the principles” in terms of general scientific understanding. This understanding may turn out to be inadequate to the task of unification, as has regularly been the case for 300 years. A good deal of discussion of these topics seems to me misguided, perhaps seriously so, for reasons such as these.

Other similarities are worth remembering. The “triumphs of chemistry” offered useful guidelines for the eventual reconstruction of physics: they provided conditions that core physics would have to meet, in some manner or other. In a similar way, discoveries about bee communication provide conditions that have to be met by some account in terms of cells. In both cases, it is a two-way street: the discoveries of physics constrain possible chemical models, as those of basic biology should constrain models of insect behavior.

There are familiar analogues in the brain and cognitive sciences: the issue of computational, algorithmic, and implementation theories emphasized particularly by David Marr, for example. Or Eric Kandel’s work on learning in marine snails, seeking “to translate into neuronal terms ideas that have been proposed at an abstract level by experimental psychologists,” and thus to show how cognitive psychology and neurobiology “may begin to converge to yield a new perspective in the study of learning” (Hawkins and Kandel 1984:380, 376). Very reasonable, though the actual course of the sciences should alert us to the possibility that the convergence may not take place because something is missing — where, we cannot know until we find out.

Questions of this kind arise at once in the study of language and the brain. By *language* I mean “human language,” and understand each particular language to be a state of a subcomponent of the brain specifically dedicated to language — as a system that is; its elements may have other functions. It seems clear that these cu-

rious brain states have computational properties: a language is a system of discrete infinity, a procedure that enumerates an infinite class of expressions, each of them a structured complex of properties of sound and meaning.

The recursive procedure is somehow implemented at the cellular level, how no one knows. That is not surprising; the answers are unknown for far simpler cases. Randy Gallistel observes that “we clearly do not understand how the nervous system computes,” even “how it carries out the small set of arithmetic and logical operations that are fundamental to any computation.” His more general view is that in all animals, learning is based on specialized mechanisms, “instincts to learn” in specific ways. These “learning mechanisms” can be regarded as “organs within the brain [that] are neural circuits whose structure enables them to perform one particular kind of computation,” as they do more or less reflexively apart from “extremely hostile environments.” Human language acquisition is instinctive in this sense, based on a specialized “language organ.” This “modular view of learning” Gallistel takes to be “the norm these days in neuroscience” (Gallistel 1997:77, 82, 86±89).

Rephrasing in terms I have sometimes used (Chomsky 1975), the “learning mechanisms” are dedicated systems  $LT(O, D)$  (*learning theories* for organism  $O$  in domain  $D$ ); among them is  $LT(\text{Human, Language})$ , the specialized “language organ,” the *faculty of language*  $FL$ . Its initial state is an expression of the genes, comparable to the initial state of the human visual system, and appears to be a common human possession to close approximation. Accordingly, a typical child will acquire any language under appropriate conditions, even under severe deficit and in “hostile environments.” The initial state changes under the triggering and shaping effect of experience, and internally determined processes of maturation, yielding later states that seem to stabilize at several stages, finally at about puberty. We can think of the initial state of  $FL$  as a device that maps experience into state  $L$  attained, hence a *language acquisition device* ( $LAD$ ). The existence of such a  $LAD$  is

guages are not only extrahuman organisms but are outside the biological world altogether, it would seem. Infants are “predisposed to learn human languages” and “are strongly biased in their choices” of “the rules underlying language,” but it is a mistake to try to determine what these predispositions are, and to seek their realization in brain mechanisms (in which case the extrahuman organisms vanish from the scene). It is worse than a mistake: to pursue the course of normal science in this case is to resort to a “magician’s trick” (Deacon 1997: chap. 4).

I have been giving quotations, because I have no idea what this means, and understanding is not helped by Deacon’s unrecognizable account of “linguistics” and of work allegedly related to it. Whatever the meaning may be, the conclusion seems to be that it is a waste of time to investigate the brain to discover the nature of human language, and that studies of language must be about the extrahuman — and apparently extrabiological — organisms that co-evolved with humans and somehow “latch on” to them, English latching on some, Japanese to others.

I do not recommend this course either; in fact could not, because I do not understand it.

Within philosophy of language and mind, and a good part of theoretical cognitive science, the consensus view also takes language to be something outside the brain: it is a property of some social organism, a “community” or a “culture” or a “nation.” Each language exists “independently of any particular speakers,” who have a “partial, and partially erroneous, grasp of the language.” The child “borrows” the language from the community, as a “consumer.” The real sound and meaning of the words of English are those of the lender and are therefore outside of my head, I may not know them, and it would be a strange accident if anyone knew them for “all of English.” I am quoting several outstanding philosophers of mind and language, but the assumptions are quite general, in one or another form.

to answer questions framed in terms of the Turing test — a serious misinterpretation of Turing’s proposals, I think, but that is another matter. For the computer model of the mind, the problems I mentioned do not arise. It also follows that nothing discovered about the brain will matter for the cognitive sciences. For example, if it is some day discovered that one interpretation of the recursive procedure can be implemented at the cellular level, and another cannot, the result will be irrelevant to the study of human language.

That does not seem to me to be a wise course.

Another approach, influential in contemporary philosophy of mind and theoretical cognitive science, is to hold that the relation of the mental to the physical is not *reducibility* but *supervenience*: any change in mental events or states entails a “physical change,” though not conversely, and there is nothing more specific to say. The preunification debates over chemistry could be rephrased in these terms: those denying the “reality” of chemistry could have held that chemical properties supervene on physical properties, but are not reducible to them. That would have been an error, for reasons already mentioned: the right physical properties had not yet been discovered. Once they were, talk of supervenience becomes irrelevant and we move toward unification. The same stance seems to me reasonable in this case.

Still another approach is outlined in a highly regarded book by neuroscientist Terrence Deacon (1997) on language and the brain. He proposes that students of language and its acquisition who are concerned with states of a genetically determined “module” of the brain have overlooked another possibility: “that the extra support for language learning,” beyond the data of experience, “is vested neither in the brain of the child nor in the brains of parents or teachers, but outside brains, in language itself.” Language and languages are extrahuman. “Languages have evolved with respect to human brains”; “The world’s languages evolved spontaneously” and have “become better and better adapted to people,” apparently the way prey and predator coevolve in the familiar cycle. Language and lan-

sometimes regarded as controversial, but it is no more so than the (equivalent) assumption that there is a dedicated *language module* that accounts for the linguistic development of an infant as distinct from that of her pet kitten (or chimpanzee, or whatever), given essentially the same experience. Even the most extreme “radical behaviorist” speculations presuppose (often tacitly) that a child can somehow distinguish linguistic materials from the rest of the confusion around it, hence postulating the existence of FL = LAD. As discussion of language acquisition becomes more substantive, it moves to assumptions about FL that are richer and more domain specific, without exception to my knowledge.

It may be useful to distinguish modularity understood in these terms from Jerry Fodor’s influential ideas (Fodor 1983). Fodorian modularity is concerned primarily with input systems. In contrast, modularity in the sense just described is concerned with cognitive systems, their initial states and states attained, and the ways these states enter into perception and action. Whether the processing (input/output) systems that access such cognitive states are modular in Fodor’s sense is a distinct question.

As Fodor puts the matter, “The perceptual system for a language comes to be viewed as containing quite an elaborate theory of the objects in its domain; perhaps a theory couched in terms of a grammar of the language” (and the same should hold for the systems of language use) (Fodor 1983:51). I would prefer a somewhat different formulation: Jones’s language L is a state of FL, and Jones’s perceptual (and production) systems access L. Theories of L (and FL) are what the linguist seeks to discover; adapting traditional terms, the linguist’s theory of Jones’s L can be called *a grammar of L*, and the theory of FL can be called *universal grammar*, but it is the linguist, not Jones, who has a theory of L and FL, a theory that is partial and partially erroneous. Jones has L, but no theory of L (except what he may believe about the language he has, beliefs that have no privileged status, any more than what Jones may believe about his visual system or problem-solving capacities).

When we look more closely, we see that more is involved here than choice of terminology, but let us put that aside. Clearly the notions of modularity are different, as are the questions raised, though they are not incompatible, except perhaps in one sense: FL and L appear to be “central systems” in Fodor’s framework, distinctive components of the central “architecture of mind,” so that the “central systems”

would not be unstructured (what Fodor calls “Quinean and isotropic”), containing only domain-neutral properties of inference, reasoning, and thought generally.

For language, this “biolinguistic” approach seems to me very sound (see Jenkins, 2000, on the state of the art). But elementary questions remain to be answered before there will be much hope of solving problems about the cellular implementation of recursive procedures, and mechanisms for using them, that appear to have evolved recently and to be isolated in the biological world in essential respects.

Problems become still more severe when we discover that there is debate, which appears to be substantive, as to how to interpret the recursive procedure. There are so-called derivational and representational interpretations, and subvarieties of each. And although on the surface the debates have the character of a debate over whether 25 is 5 squared or 5 is the square root of 25, when we look more closely we find empirical evidence that seems to support one or another view.

These are difficult and subtle questions, at the borders of inquiry, but the striking fact is that they do appear to be empirical questions. The fact is puzzling. It is far from clear what it means to say that a recursive procedure has a particular interpretation for a cognitive system, not a different interpretation formally equivalent to the first; or how such distinctions — whatever they mean — might be implemented at the cellular level. We find ourselves in a situation reminiscent of that of post-Newtonian scientists — for example, Lavoisier, who believed that “the number and nature

of elements” is “an unsolvable problem, capable of an infinity of solutions none of which probably accord with Nature.” “It seems extremely probable that we know nothing at all about ... [the] ... indivisible atoms of which matter is composed,” and never will, he thought (Lavoisier, quoted in Brock 1992:129).

Some have reacted to these problems much in the way that leading natural scientists did in the era before unification of chemistry and physics. One influential proposal is the *computer model of the mind*. According to this view, cognitive science “aims for a level of description of the mind that abstracts away from the biological realizations of cognitive structures.” It does so in principle, not because of lack of understanding we hope will be temporary, or to solve some problem for which implementation is irrelevant, or in order to explore the consequences of certain assumptions. Rather, for cognitive science it does not matter” whether one chooses an implementation in “gray matter ... , switches, or cats and mice.” Psychology is therefore not a biological science, and given the “anti-biological bias” of this approach, if we can construct automata in “our *computational* image,” performing as we do by some criterion, then “we will naturally feel that the most compelling theory of the mind is one that is general enough to apply to both them and us,” as distinct from “a biological theory of the *human* mind [which] will not apply to these machines” (Block 1990:261).

So conceived, cognitive science is nonnaturalistic, not part of the natural sciences in principle. Notice that this resembles the view of chemistry, not long ago, as a calculating device, but is far more extreme: no one proposed that “the most compelling theory of chemistry is one general enough to apply” to worlds with different physical laws than ours, but with phenomena that are similar by some criterion. One might ask why there should be such a radical departure from the practice of the sciences when we turn to the study of mind.

The account of the computer model is a fair description of much of the work in the cognitive sciences; for example, work that seeks